METHODS, SYSTEMS, AND APPARATUSES FOR QUANTITATIVE ANALYSIS OF HETEROGENEOUS BIOMARKER DISTRIBUTION

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This patent application is a continuation of U.S. application Ser. No. 16/811,688 filed on Mar. 6, 2020, which application is a continuation of U.S. application Ser. No. 15/606,122 filed on May 26, 2017, which application is a continuation of International Patent Application No. PCT/EP2015/078532 filed Dec. 3, 2015, which claims priority to and the benefit of U.S. Provisional Application No. 62/086,840, filed Dec. 3, 2014. Each of the above patent applications are hereby incorporated by reference herein in their entireties.

BACKGROUND OF THE INVENTION

Field of the Invention

[0002] The present disclosure relates to the field of automated image acquisition and analysis, particularly as applied to microscopic evaluation of diseases.

Description of Related Art

[0003] Historically available technology for evaluating tissue has permitted routine evaluation of only single genetic or protein expression/activation biomarkers in isolation. It has become clear that these valuable single biomarkers do not provide a complete picture. Sequencing technologies and biochemical measurements of protein expression and activation have focused on homogenized tissue samples and in such cases the spatial context of the expression pattern or genetic change is lost. Though the information gained through biochemical assay and sequencing technology is useful, there still remain important gaps in the information content, and an incomplete understanding of the expression and activation patterns of cells within a tumor results from the averaging of protein content from many cells.

[0004] Recent research now indicates there is important information that is missed by conventional assay technologies. One such situation is the presence of multiple genetic rearrangements or aberrations in a tumor, and the realization that if the different rearrangements occur in the same cells they can have a cooperative effect (Zong et al. 2009, Goldstein et al. 2010). Genetic inter tumor and intra tumor heterogeneity has been reported and such heterogeneity is thought to contribute to treatment failure and drug resistance in treatment (Gerlinger et al. 2012, Marusyk et al. 2012). It is therefore important to recognize not just that a tumor has multiple genetic rearrangements or deletions, but also whether these occur in the same cells, different cells or a combination of situations (Svensson, et al. 2011). Phenotypic heterogeneity and protein expression signatures have also been shown to be an important consideration in evaluating biomarkers in tumor tissue (Yap et al. 2012, Marusyk et al. 2012). Phenotypic heterogeneity may arise from genetic or epigenetic causes, and is thought to contribute to drug resistance, and relapse of cancer growth.

[0005] The ability to characterize multiple biomarkers in tissue, and to measure heterogeneity of the presence and levels of said biomarkers within and between tissues, thus will provide important information for understanding and characterizing a variety of disease states. Additionally, the ability to discern and measure the areas in tissue that have different distributions of key biomarkers may provide important information to inform development of targeted and combination therapies.

[0006] Others have attempted to analyze expression heterogeneity using different clustering methods and alternative multiplexing schemes (Gerdes et al. 2013, Qian, et al. 2010). The hierarchical clustering approach requires significant assumptions to be made. Knowing the distance between points that determines where to draw the boundary to form a new cluster is a key parameter for hierarchical clustering algorithms. Alternatively, some hierarchical algorithms (such as Ward's method (Ward 1963)) require entry of the number of clusters as a parameter. However, cut-off thresholds (distance) and number of expected clusters are both parameters that are often unknown. Additionally, some algorithms enforce assumptions about even cluster size (e.g. k-means), distance between points that are members of different clusters (hierarchical clustering) or assumptions about the expected number of clusters to be found (hierarchical clustering, k-means). Though widely used, hierarchical methods are better suited to variables measured on a discontinuous scale (e.g. +, ++, +++, ++++). For this reason, hierarchical clustering algorithms are not ideal for the requirements of expression heterogeneity analysis. Alternative density-based tools such as FLOCK (Qian, et al. 2010) have limitations in that parameters for size of hyper-regions used to calculate density and density cut-off thresholds must be estimated and entered to the algorithm to enable cluster determination.

[0007] Recently, tools such as SPADE (Qiu et al, 2012, Giesen et al. 2014) and viSNE (El-ad et al. 2013) are used for mapping hierarchical relationships between clusters of cells with high-dimensionality multiparametric expression patterns. The emphasis of such tools is to map similarity relationships between expression patterns and, in this sense, provides a different and complimentary window into the nature of multiparameter expression heterogeneity. The tools such as SPADE and viSNE were developed in the context of cytometry and place greater emphasis on mapping relationships between cell populations in high-dimensional space to visualize and classify populations outside of the context of the spatial location of expression patterns in tissue. In this sense, SPADE and viSNE represent mapping tools rather than clustering tools.

[0008] To date, we are unaware of any systems or methods that sufficiently identify clusters of heterogeneity of expression, localization, and/or activation of biomolecules within the original spatial context of cell and tissue samples.

[0009] In a multiplex slide of a tissue specimen, different nuclei and tissue structures are simultaneously stained with specific biomarker-specific stains, which can be either chromogenic or fluorescent dyes, each of which has a distinct spectral signature, in terms of spectral shape and spread. The spectral signatures of different biomarkers can be either broad or narrow spectral banded and spectrally overlap. A slide containing a specimen, for example an oncology specimen, stained with some combination of dyes is imaged using a multi-spectral imaging system. Each channel image corresponds to a spectral band. The multi-spectral image stack produced by the imaging system is therefore a mixture of the underlying component biomarker expressions, which, in